CLAIMS

1. A method comprising:

determining onsets from a music clip;

estimating tempo from an onset curve of the music clip;

determining beat candidates from the onsets;

determining from beat candidates, segments of beat sequences that are synced to an actual beat phase; and

rectifying segments of beat sequences that are out-of-sync with the actual beat phase.

2. A method as recited in claim 1, wherein the rectifying segments comprises:

building a phase tree from each segment;

searching the phase trees to determine a largest sequence of segments that share a same beat phase;

assuming that the largest sequence of segments are synced segments that follow the actual beat phase;

assuming that all segments that are not in the largest sequence of segments are out-of-sync segments; and

rectifying the out-of-sync segments.

3. A method as recited in claim 2, wherein the building comprises

determining if a subsequent segment shares the same beat phase as a current segment;

if the subsequent segment shares the same beat phase as the current segment, inserting the subsequent segment into the phase tree as a child segment of the current segment; and

iterating the previous 2 steps until all segments are processed.

4. A method as recited in claim 2, wherein the rectifying the out-of-sync segments comprises following the actual beat phase for the out-of-sync segments.

5. A method as recited in claim 1, wherein the determining segments of beat sequences comprises:

finding at least 3 continuous beat candidates having intervals of one or more tempos; and

confirming the at least 3 continuous beat candidates as actual beats synced to the actual beat phase.

6. A method as recited in claim 1, wherein the determining beat candidates comprises:

calculating a beat confidence for each onset; and

detecting beat candidates from the onsets based on the beat confidence of each onset.

7. A method as recited in claim 6, wherein the calculating comprises:

representing a rhythm pattern of the music clip with a beat pattern template; and

matching the beat pattern template along the onset curve of the music clip.

8. A method as recited in claim 6, wherein the detecting beat candidates comprises:

adaptively setting a threshold; and

comparing the beat confidence for each onset to the threshold.

9. A method as recited in claim 1, wherein the estimating tempo from an onset curve of the music clip comprises:

summing onset curves of a lowest sub-band and a highest sub-band to determine the onset curve of the music clip;

generating an auto-correlation curve from the onset curve of the music clip; and

calculating a maximum common divisor of prominent local peaks of the auto-correlation curve.

10. A method as recited in claim 9, further comprising estimating a length of a bar of the music clip.

11. A method as recited in claim 10, wherein the estimating a length comprises:

calculating the length as a maximum common divisor of three peaks in the auto-correlation curve if the three peaks are evenly spaced within the tempo of the music clip; and

if the three peaks are not evenly spaced within the tempo of the music clip, selecting the position of the maximum peak within the tempo as the length.

12. A method as recited in claim 1, wherein the determining onsets from a music clip comprises:

down-sampling the music clip into a uniform format;

dividing the music clip into a plurality of non-overlapping temporal frames;

calculating the frequency spectrum of each frame;

dividing each frame into a plurality of octave-based sub-bands;

calculating an amplitude envelope of a lowest sub-band and a highest sub-band;

detecting an onset curve from the amplitude envelope; and

determining the onsets as local maximum variances in the amplitude envelope.

13. A method as recited in claim 12, wherein the down-sampling the music clip into a uniform format comprises down-sampling the music clip to a 16 kilohertz, 16 bit, mono-channel sample.

14. A method as recited in claim 12, wherein the dividing the music clip comprises dividing the music clip into a plurality of 16 microsecond-long frames.

15. A method as recited in claim 12, wherein the calculating the frequency spectrum of each frame comprises calculating a fast Fourier transform of each frame.

16. A method as recited in claim 12, wherein the dividing each frame into a plurality of octave-based sub-bands comprises dividing each frame into 6 octave-based sub-bands.

17. A method as recited in claim 12, wherein the calculating an amplitude envelope comprises convolving the lowest sub-band and a highest sub-band with a half raise cosine Hanning window.

18. A method as recited in claim 12, wherein the detecting an onset curve from the amplitude envelope comprises calculating the variance of the amplitude envelope of each of the lowest sub-band and a highest sub-band.

19. A processor-readable medium comprising processor-executable instructions configured for:

determining beat candidates from onsets of a music clip;

estimating a tempo of the music clip;

determining from beat candidates, beat segments having sequential beats with intervals of one or more tempos;

locating synced segments that are synced to an actual beat phase;

locating out-of-sync segments that are out-of-sync with an actual beat phase; and

rectifying the out-of-sync segments.

20. A processor-readable medium as recited in claim 19, wherein the determining beat segments comprises:

finding at least 3 sequential beat candidates in a row with intervals of one or more tempos; and

confirming the at least 3 sequential beat candidates as beats that are phase-locked with the music clip.

21. A processor-readable medium as recited in claim 19, wherein the locating synced segments further comprises:

building a phase tree from each segment having sequential beat candidates;

locating segment sequences whose beat candidates share the same phase and whose combined beat candidates outnumber the combined beat candidates in other segment sequences; and

designating the located segments as synced segments.

22. A processor-readable medium as recited in claim 19, wherein the locating out-of-sync segments comprises:

finding segments that are not in a largest sequence of segments which share a same phase.

23. A processor-readable medium as recited in claim 19, wherein the rectifying comprises tracking the out-of-sync segments with the actual beat phase.

24. A processor-readable medium as recited in claim 19, comprising further processor-executable instructions configured for detecting the onsets of the music clip.

25. A processor-readable medium as recited in claim 24, wherein the detecting the onsets comprises:

down-sampling the music clip to a uniform format;

dividing the music clip into temporal frames;

calculating the spectrum of each frame;

dividing each frame into six octave-based sub-bands;

calculating an amplitude envelope from a lowest sub-band and a highest sub-band;

calculating variance of the amplitude envelope to determine an onset curve; and

extracting the onsets as local maximum variances.

26. A processor-readable medium as recited in claim 19, wherein the determining beat candidates from onsets of a music clip comprises:

calculating a confidence level for each onset; and

comparing the confidence level for each onset to a threshold.

27. A processor-readable medium as recited in claim 26, wherein the calculating comprises:

representing a rhythm pattern of the music clip with a beat pattern template; and

matching the beat pattern template along the onset curve.

28. A processor-readable medium as recited in claim 19, wherein the estimating a tempo comprises:

determining an onset curve of the music clip;

generating an auto-correlation curve from the onset curve; and

calculating a maximum common divisor of prominent local peaks of the auto-correlation curve.

29. A processor-readable medium as recited in claim 28, further comprising processor-executable instructions configured for estimating a length of a bar of the music clip.

30. A processor-readable medium as recited in claim 29, wherein the estimating a length comprises:

calculating the length as a maximum common divisor of three peaks in the auto-correlation curve if the three peaks are evenly spaced within the tempo of the music clip; and

if the three peaks are not evenly spaced within the tempo of the music clip, selecting the position of the maximum peak within the tempo as the length.

31. A computer comprising the processor-readable medium of claim 19.

32. A computer comprising:

a music clip;

a beat detection algorithm configured to detect beat candidates from onsets of the music clip and based on a tempo of the music clip; and

a rectification algorithm configured to determine segments of beat candidates that are synced with an actual beat phase and to rectify segments of beat candidates that are out-of-sync with the actual beat phase.

33. A computer as recited in claim 32, further comprising a tempo estimation algorithm configured to estimate the tempo based on an onset curve of the music clip.

34. A computer as recited in claim 33, further comprising an onset detection algorithm configured to generate the onset curve and detect the onsets from the onset curve.